



RESEARCH DEPARTMENT



REPORT

**Digital magnetic recording:
conventional saturation techniques**

No. 1972/9

RESEARCH DEPARTMENT


DIGITAL MAGNETIC RECORDING : CONVENTIONAL SATURATION TECHNIQUES

Research Department Report No. **1972/9**
UDC 681.327.13

This Report may not be reproduced in any form without the written permission of the British Broadcasting Corporation.

It uses SI units in accordance with B.S. document PD 5686.

A.H. Jones, B.Sc.
J.P. Chambers, M.A.

A handwritten signature in black ink, appearing to read 'P. Lang', written in a cursive style.

Head of Research Department

DIGITAL MAGNETIC RECORDING : CONVENTIONAL SATURATION TECHNIQUES

Section	Title	Page
	Summary	1
1.	Introduction	1
2.	Properties of saturation recording	1
	2.1. The recording process	1
	2.2. Tape loss	2
	2.3. Separation loss	2
	2.4. Losses due to replay head	2
	2.5. Azimuth loss	3
	2.6. Overall loss	3
	2.7. Pulse crowding	3
	2.8. Partial penetration	5
3.	Choice of recording code	6
	3.1. Clocked codes	6
	3.2. Self-clocking codes	6
4.	Processing of replay head output	7
5.	Current standards	8
6.	Conclusions	8
7.	References	8

DIGITAL MAGNETIC RECORDING : CONVENTIONAL SATURATION TECHNIQUES

Summary

This report gives an outline of the main features of digital magnetic recording using the commonly adopted saturation technique. Special emphasis is placed on factors limiting packing density, and results obtained during an initial experimental investigation are quoted. Some possible coding and signal processing techniques are also described.

1. Introduction

The advantages of processing and distributing television and sound signals in digital form are well known.^{1,2} These advantages would be particularly valuable in the recording field. A survey of current developments in high-density digital recording³ has therefore been carried out, and indicated that magnetic recorders of conventional type could be used to handle the data rate required for digital television, albeit at a consumption of medium greatly in excess of that required for conventional analogue recording. Such a digital recorder could therefore provide a useful basis for experimental work, although ultimately a different recording technology, and therefore recording medium, will probably be used.

An experimental digital sound recorder⁴ has been constructed in order to furnish experience in the signal processing required for digital television recording. Concurrently, some basic studies of the techniques used in conventional digital magnetic recording and replay were made. These were confined to saturation recording;* similar work on non-saturation recording is in progress.

The magnetic recording process is by no means completely understood, and there are several different detailed theories receiving support. Nevertheless, this report gives an account of the generally accepted mechanism of the conventional saturation record and replay processes.

The limiting packing density of information on tape depends on the characteristics of the tape, heads, transport

and signal processing. However, with the equipment available for the present study, it was difficult to obtain consistent results when working near the limiting packing density, and this was itself considerably below the limit claimed for the best components currently available. Thus it is not possible to infer an ultimate packing density from the results of this study. The results were, however, never inconsistent with prevailing theories and current 'rules of thumb', and serve to illustrate the basic processes involved. Some consideration is given to means of matching the limitations of the recording system to the tolerances of the signal to be recorded.

2. Properties of saturation recording

The characteristics of a saturation recording system can, in part, be inferred from the response of the system to the transitions in the two-level digital input signal, these corresponding ideally to transitions between extremes of saturation on the tape. Although the overall process is not linear, and the principle of superposition cannot rigorously apply, the superposition of the replayed pulses corresponding to individual transitions is often used to provide a 'model' of the system, and the results obtained are reasonably consistent with those obtained in practice. Factors influencing the width of the replayed pulses, and therefore the maximum information packing density, will now be considered.

2.1. The recording process

The recording head is driven with a bipolar current waveform corresponding to the two-level digital signal; the head inductance is usually such that the rise-times of the current waveform are insignificant. The flux pattern set up around the record head gap gives rise to a level of magnetisation which varies with distance into the tape coating. The state of tape magnetisation finally reached is determined by the behaviour of the changing field from the head and it is

* Some confusion exists in the literature as to the usage of the term 'saturation recording'. In this report the expression implies that the magnetic coating is taken into saturation through at least part of its thickness. This contrasts with techniques using h.f. bias, in which none of the tape is left in a fully saturated condition.

generally held that the final state of magnetisation is reached at the region where the field is about equal to the coercivity of the tape coating. This region is usually offset from the trailing edge of the recording gap by a distance comparable to the gap length, and it has a length (measured along the tape) which depends on the hysteresis characteristics of the tape and the magnitude of the recording current,⁵ being typically of the order of one-third the tape coating thickness if the current is set to record right through the coating.⁶

As the recording zone has a finite size there is an associated aperture loss which restricts the sharpness of the recorded transition.

2.2. Tape loss

There is a limit, set by the magnetic properties of the coating, to the maximum gradient of magnetisation which can be maintained within the tape coating. When saturation recording is used, this limitation outweighs that caused by the finite width of the zone in which recording takes place. The theory has been developed^{7,8} on the assumption that the transition in the magnetisation remaining on the tape may be expressed as

$$M(x) = \frac{2M_r}{\pi} \tan^{-1} \frac{x}{a}$$

where M_r — remanent magnetisation after saturation
 x — distance along tape from centre of transition
 a — 'character distance' given by

$$a = \frac{B_r c}{2\pi H_c}$$

where c — coating thickness

B_r — coating remanence (Gauss) ($= 4\pi M_r$)

H_c — coating coercivity (Oersted)

This function is convenient for analysis and gives results in reasonable agreement with the observed waveforms from a playback head, allowing for other losses occurring during playback. Neglecting these additional losses, the replayed pulse, is determined by $(dM(x))/(dx)$. It has a half-width which is usually referred to in the literature as P_{50} and is equal to $2a$.

Fig. 1 shows the idealised transition $M(x)$ and the corresponding replayed pulse $(dM(x))/(dx)$.

2.3. Separation loss

In practice, there is always a significant separation between the magnetised coating (which has a finite thickness) and the replay head (whose poles do not reside on the surface). This results in further losses, the signal components of wavelength λ along the tape being attenuated by a factor $e^{-2\pi\lambda/d}$, where d is the separation between the gap and the tape-coating layer giving rise to the signal. This loss is often quoted as 55 dB per wavelength of separation. It follows that, at high frequencies (i.e. where the recorded wavelength is comparable to the coating

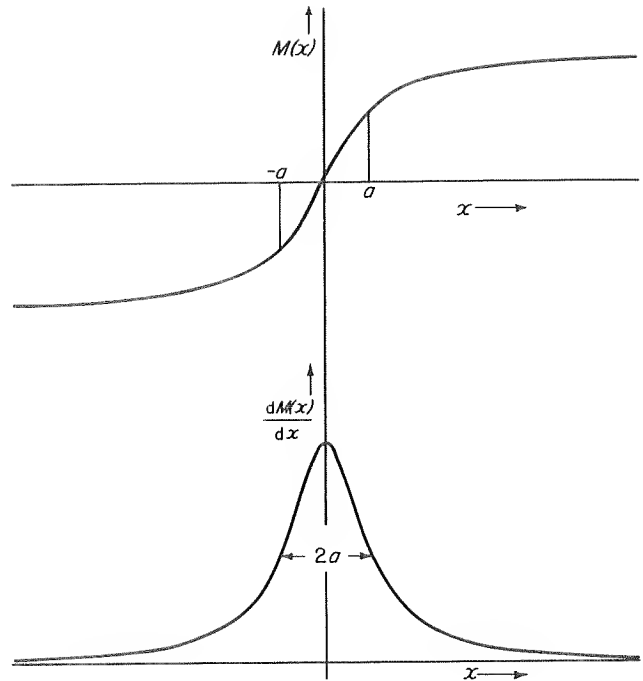


Fig. 1 - Idealised form of transition in tape magnetisation and corresponding replayed pulses assuming no further losses

thickness), most of the output is contributed by the outer layers of the coating. For example 75% of the output produced from a fully saturated recording comes from the outer 0.22λ of coating thickness.

2.4. Losses due to replay head

Magnetic losses in the replay head can usually be made of little significance. The choice of material for the core and pole tips, and the dimensions and winding configuration used can be optimised for any normal requirement.

However, the replay head gap must be of finite length, measured along the tape, in order to divert flux through the core of the head. This finite gap length leads to an aperture loss (K_1) of the form

$$K_1 = \frac{\sin \frac{\pi l}{\lambda}}{\frac{\pi l}{\lambda}}$$

where l = gap length

λ = recorded wavelength

Because the core material has a finite permeability, the performance of a head is usually inferior to that predicted by taking l as equal to the physical gap length, and it is usual to define as the 'effective gap length' the recorded wavelength that corresponds to the frequency at which the first null in the head output occurs when a sinewave of increasing frequency is applied to the recording head. For a well-made replay head, the effective gap length is about 30% greater than the physical gap length.

2.5. Azimuth loss

If there is an angular misalignment α between the gaps used for recording and replay, a further aperture loss (K_2) occurs, given by

$$K_2 = \frac{\sin \frac{w \tan \alpha}{\lambda}}{\frac{w \tan \alpha}{\lambda}}$$

where w = recorded track width.

This is an idealised result, but sufficient to indicate the tolerable azimuth errors. This loss is, of course, less critical when narrower tracks are used.

2.6. Overall loss

An overall formula compounding tape loss, thickness loss and separation loss^{7,8} gives the half-amplitude width P_{s0} of an isolated pulse (equivalent length of tape) as

$$P_{s0} = 2 \sqrt{(a+d)(a+d+c) + \left(\frac{l}{2}\right)^2}$$

a — as defined in 2.2

d — separation of head from coating

c — coating thickness

l — replay effective gap length

This formula enables the benefits obtained by improving each parameter in a given situation to be assessed. It is of particular use in disc recording systems where the head separation is well defined.

Fig. 2 shows a typical example of replayed pulses corresponding to a low-frequency square wave recorded at a current sufficient to saturate the tape throughout the coating. A close up of one of the pulses is given in Fig. 3; its P_{s0} was measured as 205 $\mu\text{in.}$ (5.2 $\mu\text{m.}$).

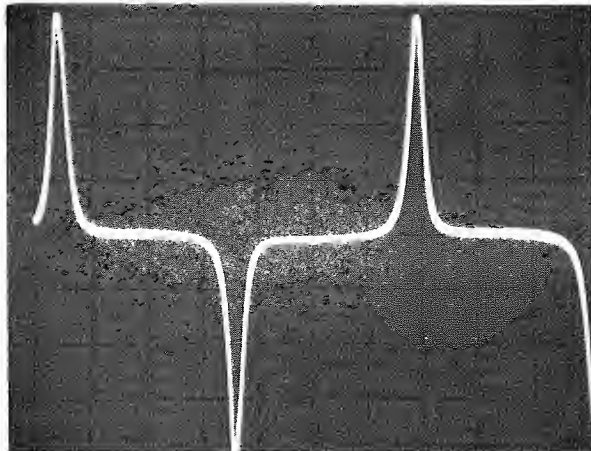


Fig. 2 - Replayed waveform from recorded low-frequency square wave

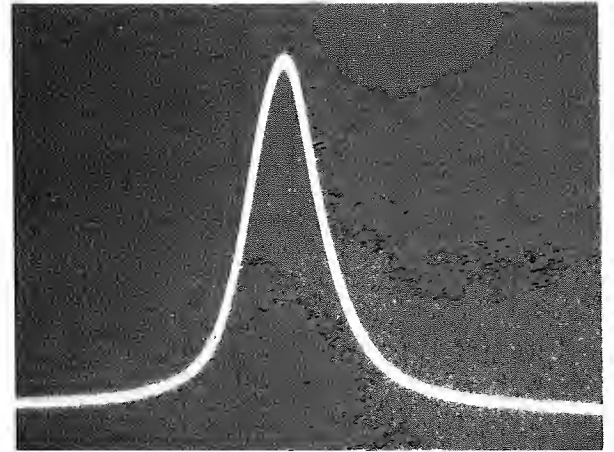


Fig. 3 - Expanded version of one of the pulses shown in Fig. 2

The first extinction frequency for the replay heads used in this case was estimated at about 400 kHz at 15 ips (38 cm/sec), this suggests that $l = 38 \mu\text{in.}$ (1 $\mu\text{m.}$). Relevant tape parameters (3M type 971) were: $c = 110 \mu\text{in.}$ (2.8 $\mu\text{m.}$), $B_r = 1500$ gauss, $H_c = 530$ oersteds. Therefore $a = 50 \mu\text{in.}$ (1.27 $\mu\text{m.}$).

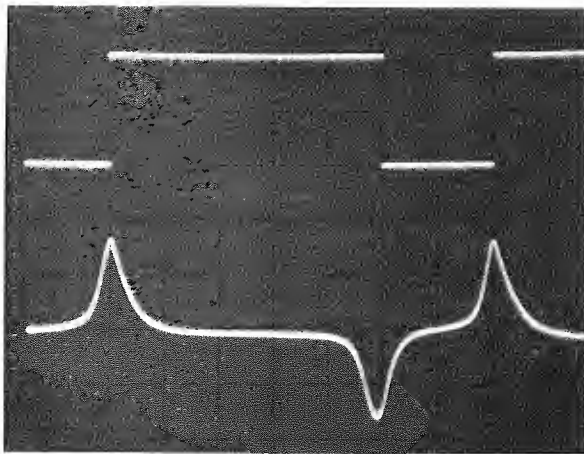
Estimating d as 15 $\mu\text{in.}$ (0.38 $\mu\text{m.}$) and substituting into the above formula one obtains $P_{s0} = 168 \mu\text{in.}$ (4.3 $\mu\text{m.}$).

2.7. Pulse crowding

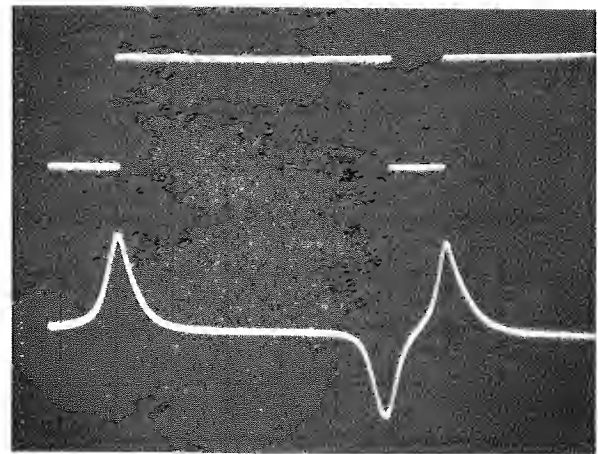
At high packing densities the replayed pulses are no longer isolated; their overlapping and consequent interaction can impair the detection process and limit the usable packing density. The usual type of detection process assumes that the peaks of the head output waveform indicate the location of the recorded transitions, and the effect of pulse interaction is to separate those peaks which are close together. This is clearly shown in Fig. 4; the closer the second and third transitions in the recorded signal become, the less accurately are their positions conveyed by the replayed pulses. Fig. 4 also indicates that for very close transitions the principle of superposition no longer applies.

This 'pulse-crowding' effect can be reduced if one anticipates it by modifying the record waveform; the transitions that will be separated are brought closer together in the recorded signal. Such predistortion is relatively easily carried out by referring to the timings of the previous and later transitions. The improvement obtained is relatively small, however, in practice it permits an increase in packing density of only some 15%.

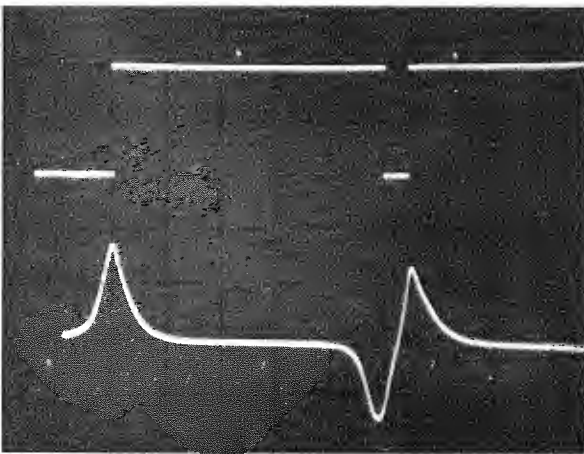
The limit imposed by pulse-crowding is often quoted in terms of the response characteristic obtained by plotting the peak-to-peak magnitude of the replayed signal against the frequency of a recorded square-wave signal. The curve obtained stays at maximum level until the replayed pulses interact, and then begins to fall. When the -3 dB point is



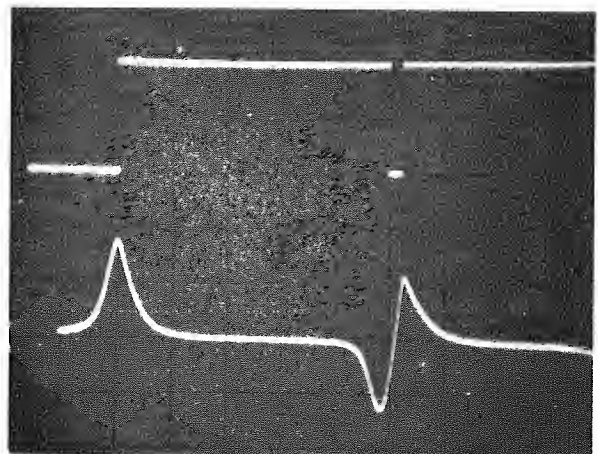
(a)



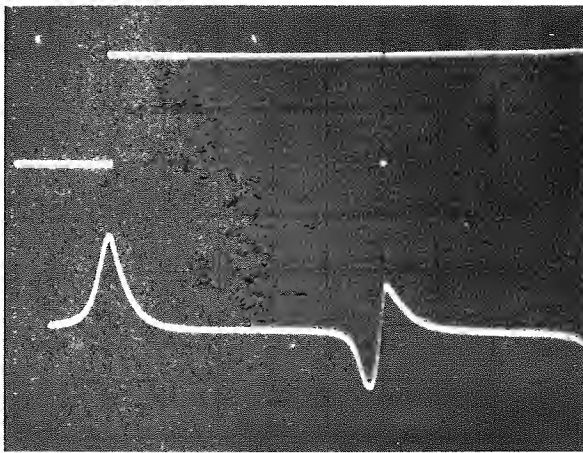
(b)



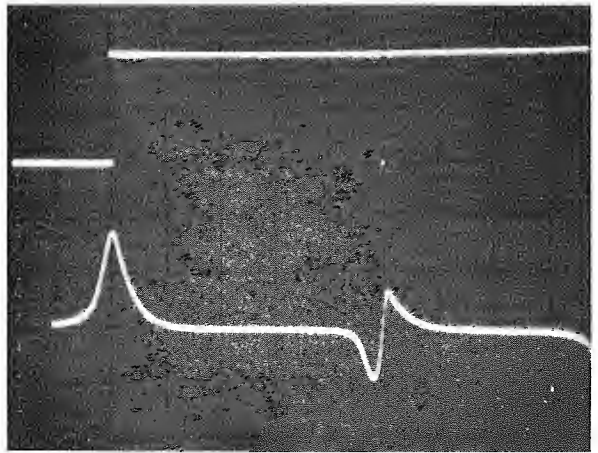
(c)



(d)

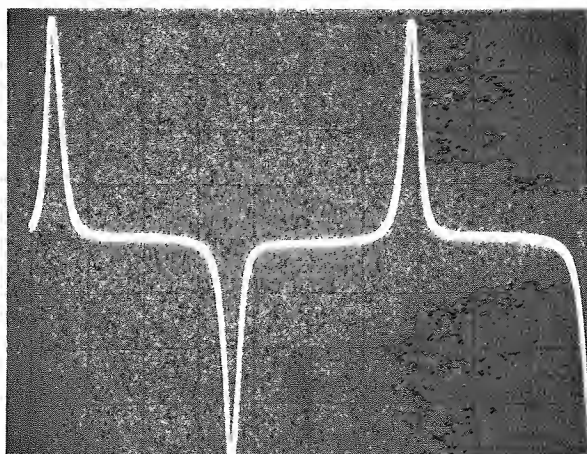


(e)

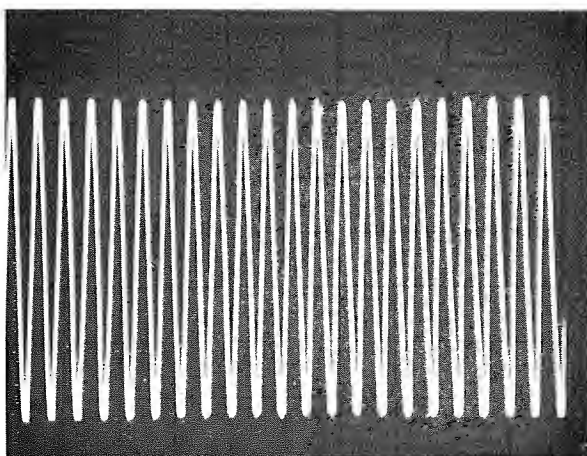


(f)

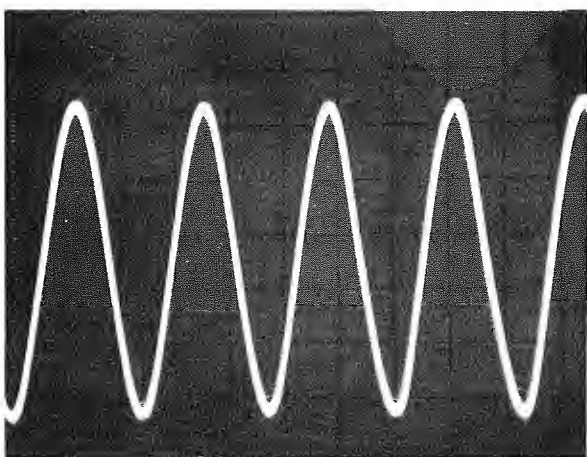
Fig. 4 - Effect on replayed waveform of inter-pulse crosstalk



(a)



(b)



(c)

Fig. 5

- (a) Replayed waveform corresponding to low-frequency square-wave (from Fig. 1)
 (b) Frequency increased until pk/pk amplitude of replayed signal falls by 3 dB
 (c) Expanded version of (b)

reached, the output is roughly sinusoidal, as Fig. 5 shows. The distance between successive flux transitions is then roughly equal to P_{50} , and the maximum packing density for relatively simple recording systems demanding an extremely low error rate, e.g. those for computer recorders, has been reached. Where more sophisticated techniques are used and the error rates need not be quite so low, it is found that a higher packing density, e.g. that given by the -5 dB point,* can be used.

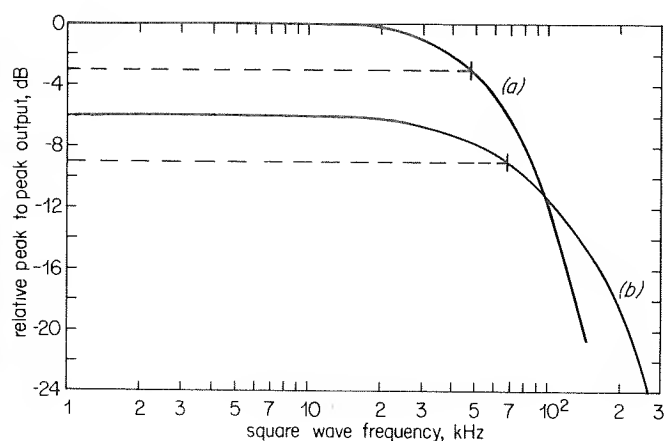


Fig. 6 - Response/frequency characteristics, square-wave recording

(a) tape fully saturated (b) tape partially penetrated

Fig. 6 shows at (a) the square wave response of the typical system quoted above. The -3 dB point is at 48 kHz, from which $P_{50} = 156 \mu\text{in.}$ ($4.0 \mu\text{m.}$). One flux reversal every $156 \mu\text{in.}$ suggests a packing density of 6400 bits per inch (one flux transition per bit).

2.8. Partial penetration

The formulae given in Sections 2.2 and 2.6 make it clear that coating thickness is an important parameter determining the width of the replayed pulses. One way of reducing the effective coating thickness and thereby increasing the packing density is to reduce the recording current. The field around the recording head gap is thus reduced, and the recording penetrates only a little way into the coating. The effect is approximately equivalent to that produced by having a thinner coating.

This is illustrated in Fig. 5(b), which shows the square-wave response of the above system plotted after the recording current had been reduced so as to give half the peak-to-peak output at low frequencies. The -3 dB point has now increased in frequency to 67 kHz, and the corresponding packing density is 9k bits per inch.

This reduction in the amplitude of the replayed signal decreases the noise margin, and the increased packing

* Some further improvement may be obtained by including equalising circuits to offset the falling response at high frequencies. Such equalisation reduces the noise margin, however.

density is therefore achieved at the cost of an increase in error rate. Further, partial penetration appears to increase the incidence of signal dropouts; the degree to which it can be used is therefore dependent on the overall system requirements.

3. Choice of recording code

The foregoing description has indicated some of the factors that limit packing density. In practical applications of digital recording there is always a complicated interplay between these and other factors such as mechanical stability, permissible error rates, instrumental limitations, etc., and it is in the light of all of these considerations that one looks for the signal coding and decoding technique which will make the best use of the recording medium. The properties of some of the coding techniques currently in use will now be considered.

The recorded information may be considered to be contained either in the direction of magnetisation existing in successive bit-cells along the length of the tape, or in the transitions occurring within or between bit-cells. Whether the detection process takes account of the direction of magnetisation or the existence of transitions, or both, it is important to have available within the replay circuits a clock wave which is accurately locked to the data being read from the tape. Moreover, the data retrieved from the tape is bound to be subject to timing errors caused by mechanical imprecision, etc., and therefore the clock wave must be derived from tape signals. A separate clock track may be used for this purpose or additional transitions may be introduced into the data stream so as to permit the recovery of a clock from the data itself. Several examples of clocked and self-clocking codes are given in Fig. 7.

3.1. Clocked codes

The simplest code for use with the separate clock track is the simple 2-level NRZ (non-return to zero) waveform, Fig. 7(a). Because a signal coded in this form can, when the data is unchanging, maintain one level for an indefinite time, the head output cannot be fully equalised by integration to recover the recorded magnetisation pattern. The transitions can however be identified and used to drive a bistable circuit which can then be sampled by the signal derived from the clock track.

Another clocked code having similar properties, NRZ 1 (Fig. 7(b)) records a transition when a 1 appears at the input and no transition for an 0. As Fig. 7(b) indicates, the same information is conveyed if the polarity of the signal is reversed — a convenient property in practice.

Clocked codes usually necessitate less processing for detection as compared with self-clocking codes and are therefore often used when the recording contains many parallel tracks. Because the clock wave is derived from a separate track, however, it is susceptible to errors in head alignment; it is possible to allow for static timing errors between the clock and data tracks by using preset adjustable delays. An improvement might be achieved if circuits are

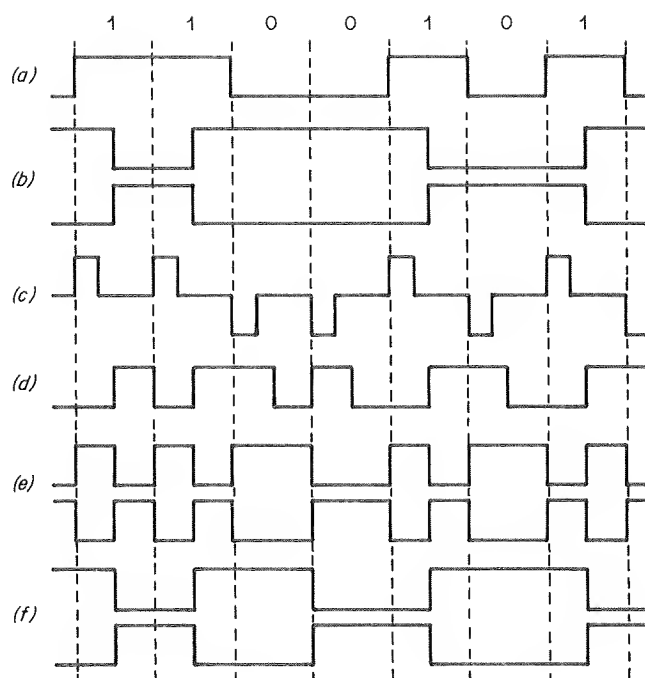


Fig. 7 - Some of the common recording codes
(a) NRZ (b) NRZ 1 (c) RZ (d) Biphas L
(e) Biphas M (f) DM

introduced to generate a clock wave having a frequency determined by the clock track but whose phase is determined by the timing of transitions in the data tracks.

3.2. Self-clocking codes

When timing information is to be derived from the data track alone, an extrapolation process is required. On the basis of the timing of the most recently occurring transitions, an estimate can be made of the possible timings of subsequent transitions. When the subsequent transitions occur, their intended timings are then inferred from the predicted possibilities. The prediction process then continues using the new transitions as its input data.

The RZ (return-to-zero) code illustrated in Fig. 7(c) is a simple type of self-clocking code which is literally self-clocking in that the detection process does not need access to a regenerated clock wave. This code has many disadvantages, however; the recorded signal is a three level waveform needing a wide bandwidth, and both packing density and noise immunity are poor.

The most popular self-clocking code is Biphas, one example of which, known as Biphas L, is shown in Fig. 7(d). The value of each bit is indicated by the direction (+ve or -ve going) of the transition occurring at the centre of the bit cell. There are two possible transition spacings, $\frac{1}{2}$ and 1 bit cell, and a twice bit-rate clock wave is readily derived from the data. The signal may be decoded by comparing its values during the first and second halves of each bit cell. There is always a transition in the centre of a bit cell; this permits the cell boundaries to be defined.

Biphase M, Fig. 7(e) is an alternative in which transitions are always present at bit-cell boundaries and additionally at the centre of those bit-cells in which a 1 is being conveyed. As Fig. 7(e) shows, this code also conveys the same information when the polarity of the signal is reversed; the arrangement is relatively easily instrumented.

Delay modulation⁹ or Miller Code, shown in Fig. 7(f) is a more recent alternative. Here, transitions are introduced at the centre of each 1 cell and between adjacent 0 cells. Inspection of Fig. 7(f) shows that this is equivalent to using alternate Biphase transitions only. It might be expected that with fewer transitions, a higher packing density could be achieved (and this conclusion is supported by the observation that the spectrum of delay-modulated data does not extend as far as that of Biphase-modulated data). It should be borne in mind, however, that the transition spacings now need to be categorised into three possibilities instead of two, and inter-pulse cross-talk is therefore more troublesome. So far as bit packing-density is concerned, therefore, it may be that neither system has a clear advantage over the other. The error curve shown in

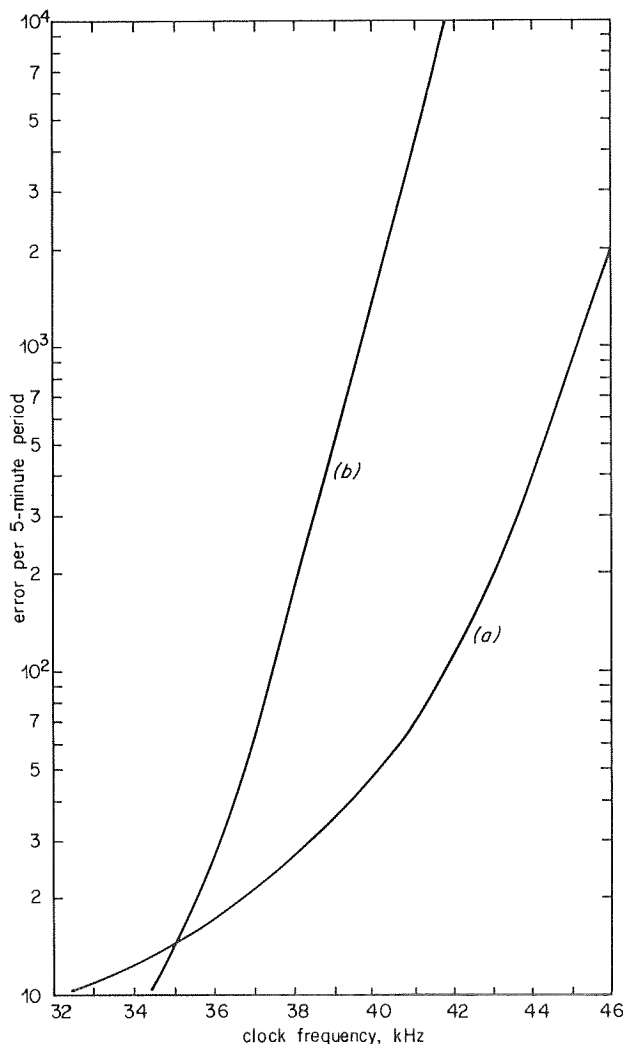


Fig. 8 - Typical error count vs. clock frequency curves obtained using:

(a) Biphase L modulation (b) Delay modulation

Fig. 8 indicates that under typical experiment conditions* delay modulation failed relatively suddenly at a bit packing density at which biphase was already beginning to contribute errors. This rapid failure occurs because there is an ambiguity in clock phasing until a particular sequence of transitions corresponding to three successive bits having values 1, 0 and 1 arrives; at packing densities for which the system is about to fail it is possible for the correct clock/data phasing to be lost, leading to a burst of errors, until it is re-established by another 1 0 1 sequence.

Fig. 7(f) indicates that delay modulation is a coding system (like Biphase M) for which the output signal is not affected when the input to the detector is phase reversed.

4. Processing of replay head output

The terminal e.m.f. from the replay head windings is essentially the derivative of the flux distribution along the tape. This signal must first be brought to a usable level by low noise amplification. A number of means of recovering the recorded information, given by the timings of the transitions in the recorded waveform, are then available.

The signal may for example be equalised, as with a conventional analogue recorder, to compensate for the response/frequency characteristic of the system, the main component of the equalisation being an integration (bass boost). The resulting signal will then be an approximation to the recorded waveform and may be restored to a two-level form by hard limiting. If the recorded waveform has a considerable low-frequency content (wide possible spacing of transitions), this approach fails due to noise limitations when the low frequencies are emphasised.

As an alternative method, the signal may be differentiated in order to locate the peaks of the replayed waveform (corresponding to the transitions in the recorded signal) as zero-crossings of the derivative. This method also fails when the transitions are widely spaced because the derivative of the signal between the replayed pulses returns to zero. A self-clocking code must therefore be used, and it is necessary to ensure that even when the spacing between successive transitions is at its greatest there is still an overlap between pulses, i.e. the replayed waveform must have zero gradient only at its peaks. This method may be conveniently instrumented, and was used to obtain the results given in Fig. 8.

A third and perhaps the best approach to the regeneration of the recorded signal would be to use a matched filter or correlation technique to form best estimates of the positions of the replayed pulses. The advantage of this process is that it takes account of the whole, and not just the peaks, of the replayed signal.

Experiments are being conducted on a number of different coding and processing systems. One aspect which particularly influences the recovery of the data is the provision of a reliable regenerated clock wave; the optimum

* Fig. 8 was obtained using heads having parameters inferior to those quoted in Section 2.6.

characteristics of a phase-locked loop clock regenerator are therefore being examined. Results will be given in a later report.

5. Current standards

High level saturation recording is commonly used in computer recorders where packing density is of secondary importance to freedom from errors.

The standard format for computer tape systems is 7 or 9 tracks of about 1.1 mm (28 thou) width laid longitudinally on 12.7 mm ($\frac{1}{2}$ in.) tape. Using replay head gap lengths of $6.3\ \mu\text{m}$ (250 $\mu\text{in.}$) the packing density achieved is 32 flux transitions per mm (800 FRPI), corresponding to 800 bits per inch using the NRZ1 code. Using head gaps of $2.4\ \mu\text{m}$ (95 $\mu\text{in.}$) 125 flux transitions per mm (3200 FRPI) is achieved, corresponding to 1600 bits per inch using biphase encoding.

According to the requirement the tape speed may be one of several possibilities up to 120 ips.

The tape is a conventional oxide tape with a coating thickness of about $10\ \mu\text{m}$ (400 $\mu\text{in.}$). The entire length is tested during manufacture and its surface finish is certified. The tape is expected to withstand the rapid accelerations of a stop-start transport.

The above standard is framed to give good margins against noise and mistiming effects to ensure low inherent error rates (better than 1 in 10^7). The error rate is often further improved by recording redundant information as a check on the data (for example, the simple parity check).

Digital instrumentation recorders do not generally involve rapid starting and stopping, and a higher error rate can usually be tolerated. Tape widths of $\frac{1}{2}$ in. and 1 in. are common, and thinner magnetic coatings and backing materials are used. The transports are usually capable of a wide range of speeds between 1 7/8 and 120 ips. Track density has until recently been 14 per inch but 28 tracks per inch is becoming common, and 42 tracks per inch is now in use. Track density is chiefly limited by difficulties in head design, though higher densities make greater demands on tape guidance. Within each track, packing densities range from 1000 to 20,000 bits per inch, though some experimental work has been done at over 30,000 bits per inch. 20,000 bits per inch is thought to give error rates acceptable for experimental digital television recording.

6. Conclusions

In high-density digital magnetic recording both heads and tape limit the maximum usable packing density. The limitations imposed by the tape can be reduced by partial penetration, but only if an increase in error rate can be tolerated. These factors together with the format of the data and a number of mechanical and instrumental considerations determine the signal coding and processing techniques to be adopted for each application.

7. References

1. Digital methods in television. BBC Research Department Report No. EL-12, Serial No. 1967/55.
2. Pulse-code modulation for high-quality sound-signal distribution: appraisal of system requirements. BBC Research Department Report No. EL-10, Serial No. 1967/50.
3. Possible techniques for the recording of digital television signals. BBC Research Department Report No. 1969/42.
4. Proposed form of an experimental digital sound recorder. BBC Research Department Technical Memorandum No. PH-1015.
5. McCARY, R.O. 1970. Saturation magnetic recording process. *IEEE Trans. Magnet.*, 1971, **MAG-7**, 1, pp. 4 – 16.
6. CAMRAS, M. 1964. An X-field micro-gap head for high density magnetic recording. *IEEE Trans. Broadcast & Telev. Receiv.*, 1964, **BTR-10**, 1, pp. 93 – 104.
7. BONYHARD, P.I., DAVIES, A.V. and MIDDLETON, B.K. 1966. A theory of digital magnetic recording on metallic films. *IEEE Trans. Magnet.*, 1966, **MAG-2**, 1, pp. 1 – 5.
8. SPELIOTIS, D.E. and MORRISON, J.R. 1966. A theoretical analysis of saturation magnetic recording. *IBM Jnl*, **10**, 2, pp. 233 – 243.
9. HECHT, M. and GUIDA, A. 1969. Delay modulation. *Proc. IEEE*, 1969, **57**, 7, pp. 1314 – 6.